Supplementary Information for:

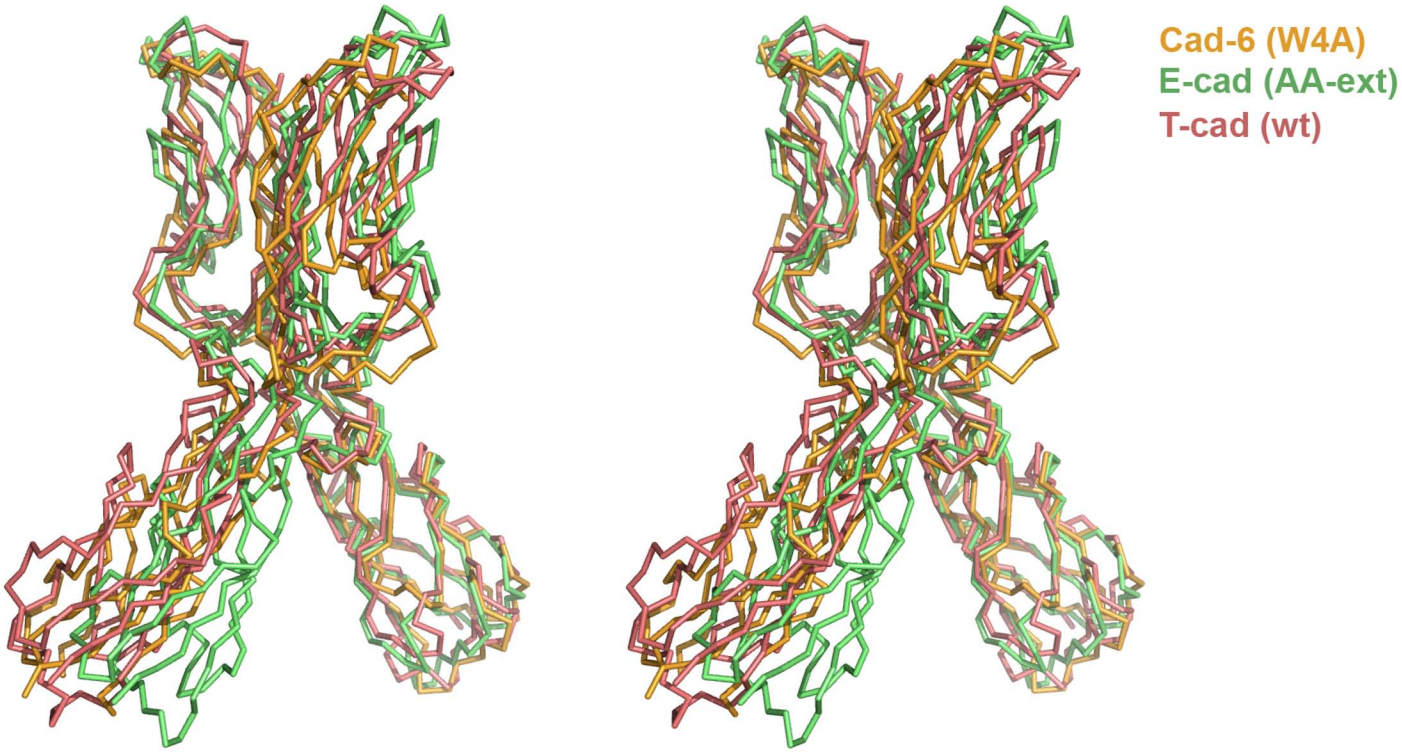
Two-step adhesive binding by classical cadherins

Oliver J. Harrison, Fabiana Bahna, Phinikoula S. Katsamba, Xiangshu Jin, Julia Brasch, Jeremie Vendome, Goran Ahlsen, Kilpatrick J. Carroll, Stephen R. Price, Barry Honig*, and Lawrence Shapiro*

* To whom correspondence should be addressed:

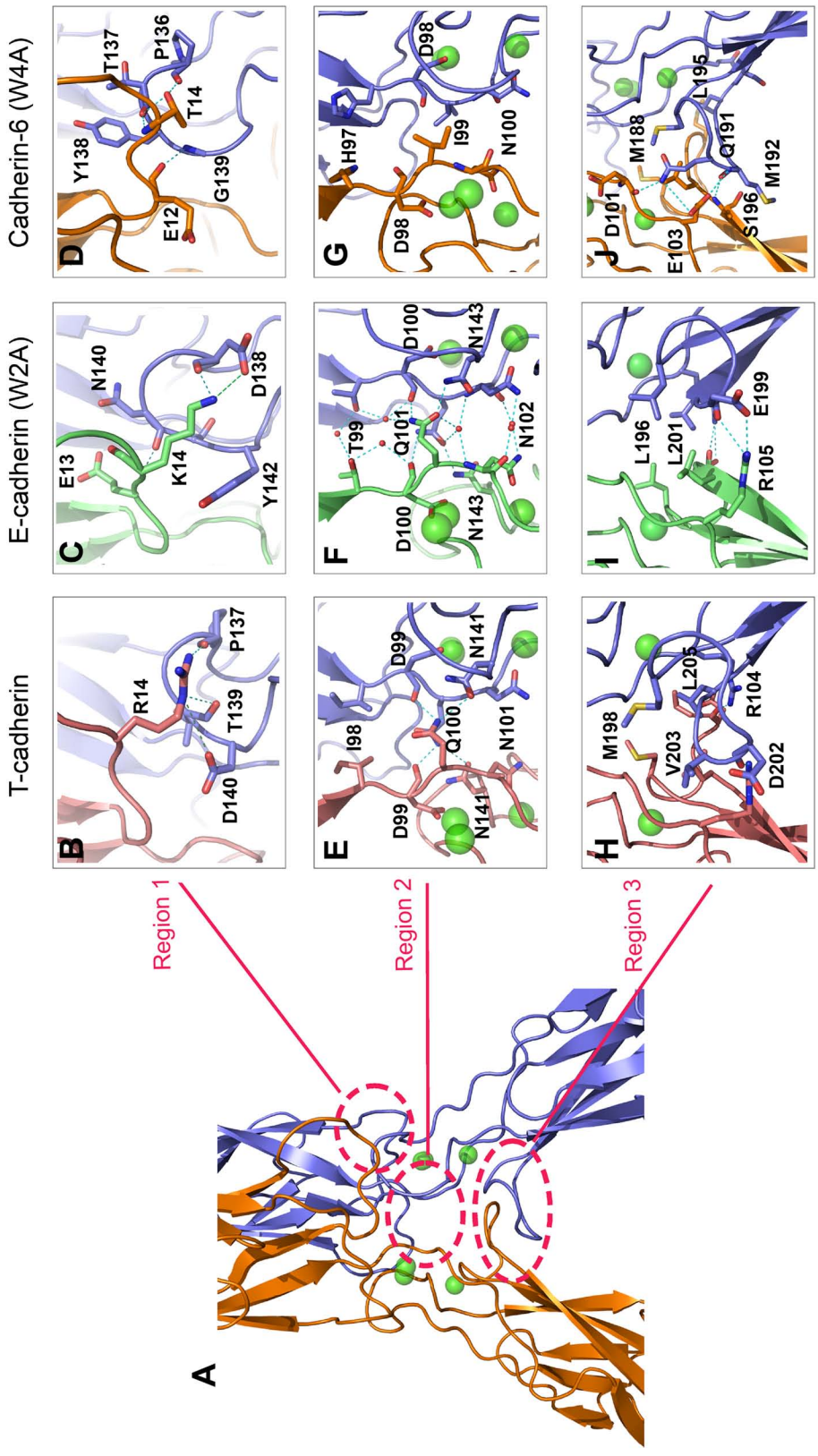
LS: LSS8@columbia.edu

BH: BH6@columbia.edu

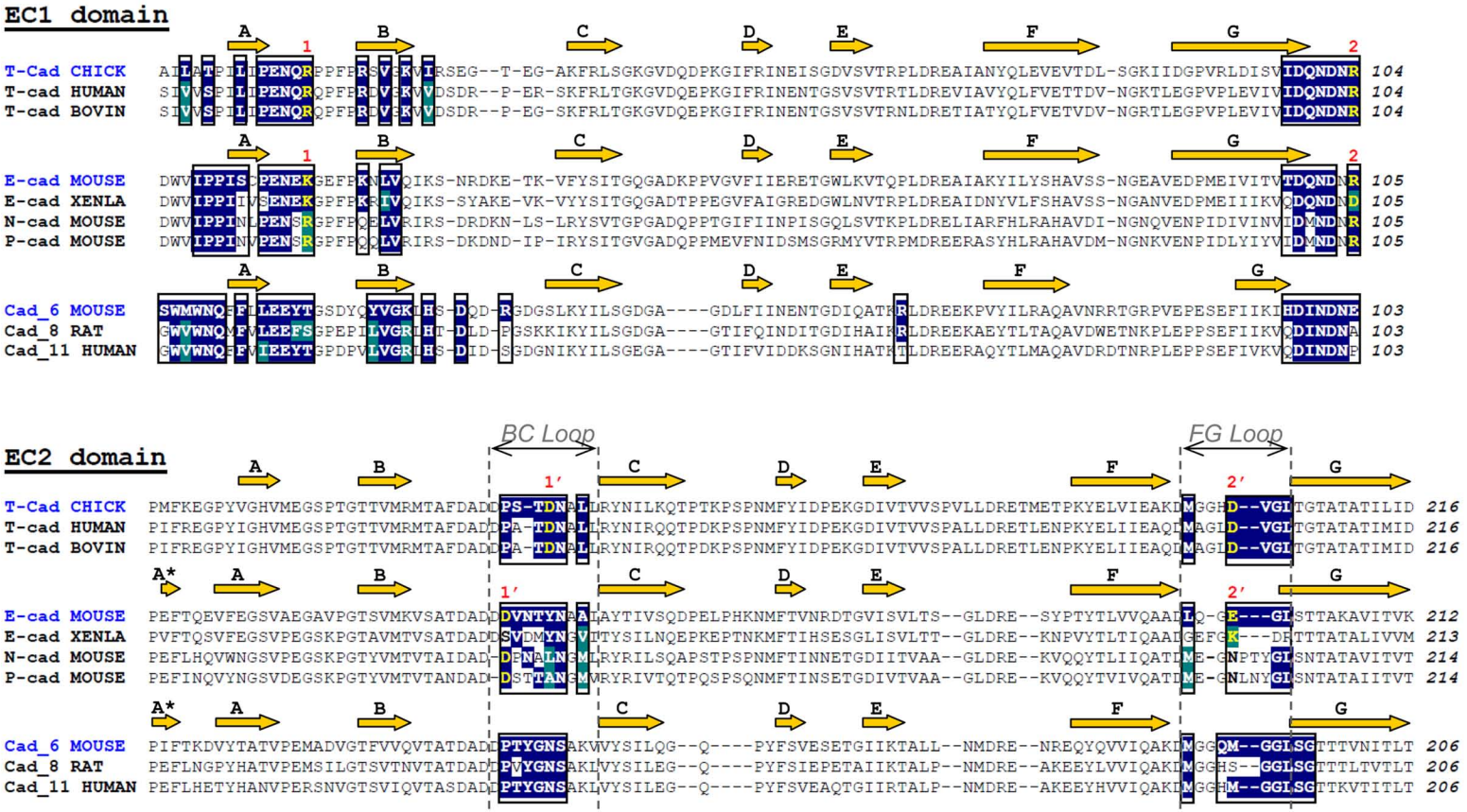


Supplementary Figure 1: Superposition of T-cadherin, mutant E-cadherin and mutant cadherin-6 X-dimers. Stereo image showing superposed α-carbon traces of X-dimers of T-cadherin (red), E-cadherin EC1-2 AA-extension (green) and cadherin-6 EC1-2 W4A (orange). Superposition is based on all residues of one of the two protomers (shown in background) to highlight differences in dimer angle.

Supplementary Figure 2

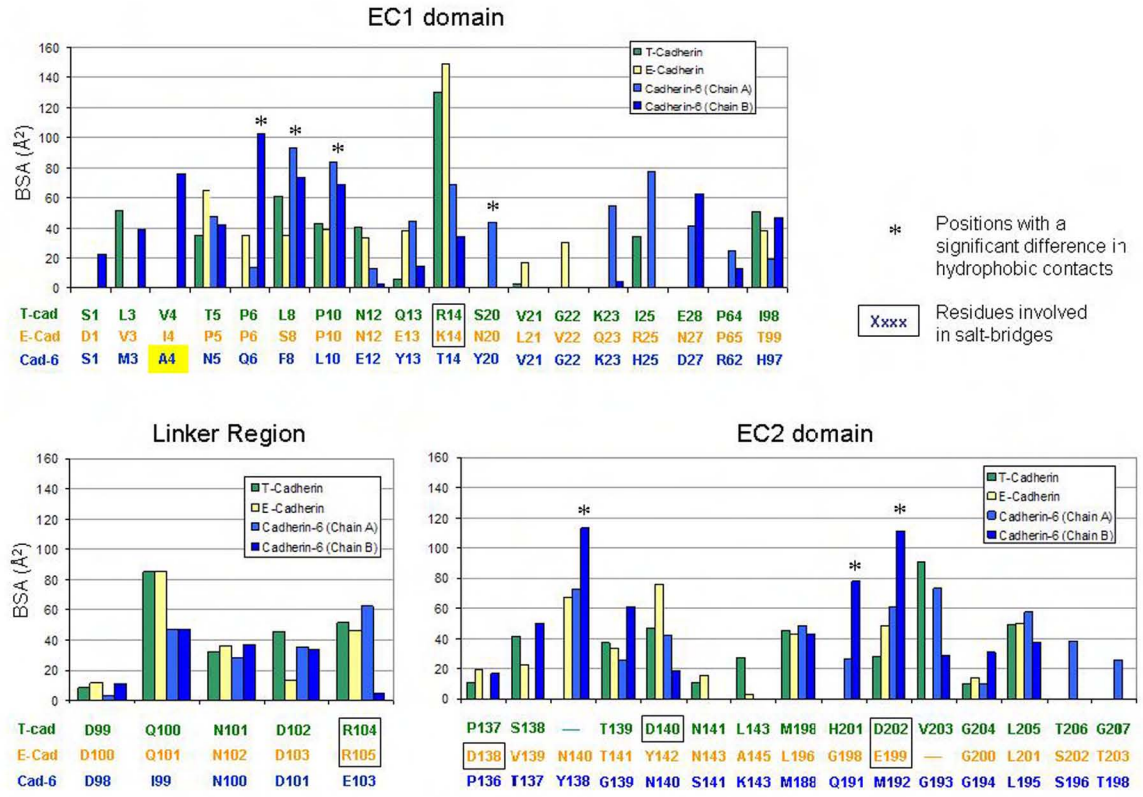


Supplementary Figure 2: X-dimer interfaces in T-cadherin, mutant E-cadherin and mutant cadherin-6. Three regions of the X-dimer interface, circled in (A), are shown in detail for chicken T-cadherin EC1-2 from Ciatto *et al* 2009 (B, E, H), W2A strand dimer mutant E-cadherin EC1-2 (C, F, I) and W4A strand dimer mutant cadherin-6 EC1-2 (D, G, J). See supplementary text for explanation. Side chains of residues involved in intermolecular interactions are shown; hydrogen bonds and ionic bonds are indicated by blue and green dashed lines respectively. Calcium ions are shown in green, water molecules in red.

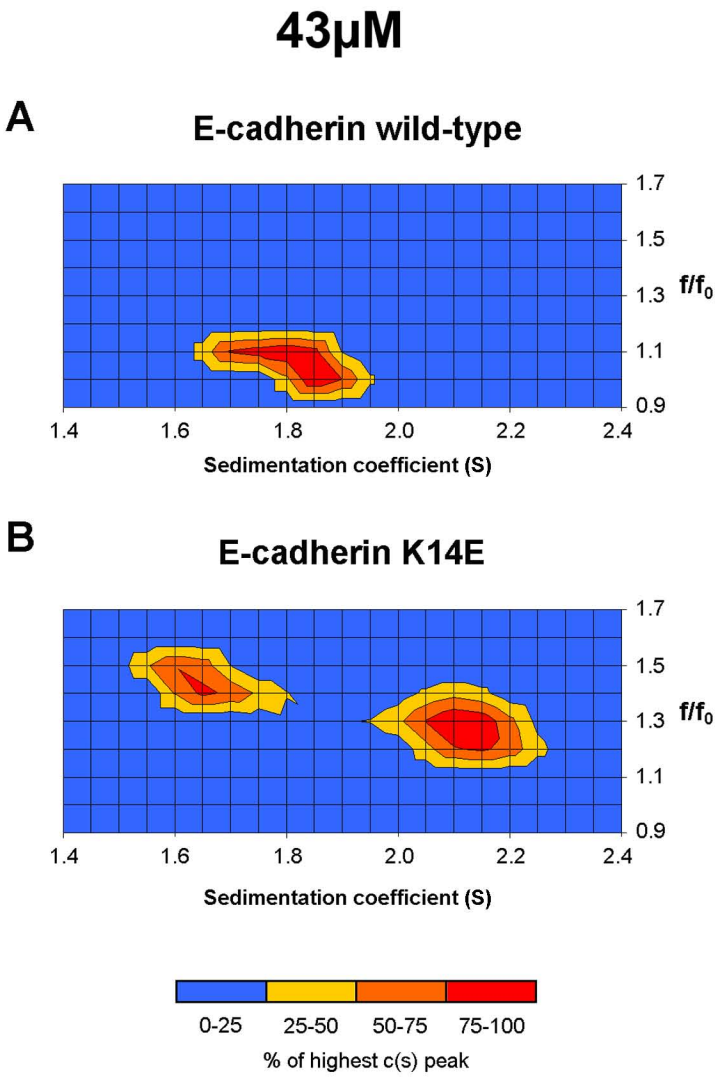


Supplementary Figure 3

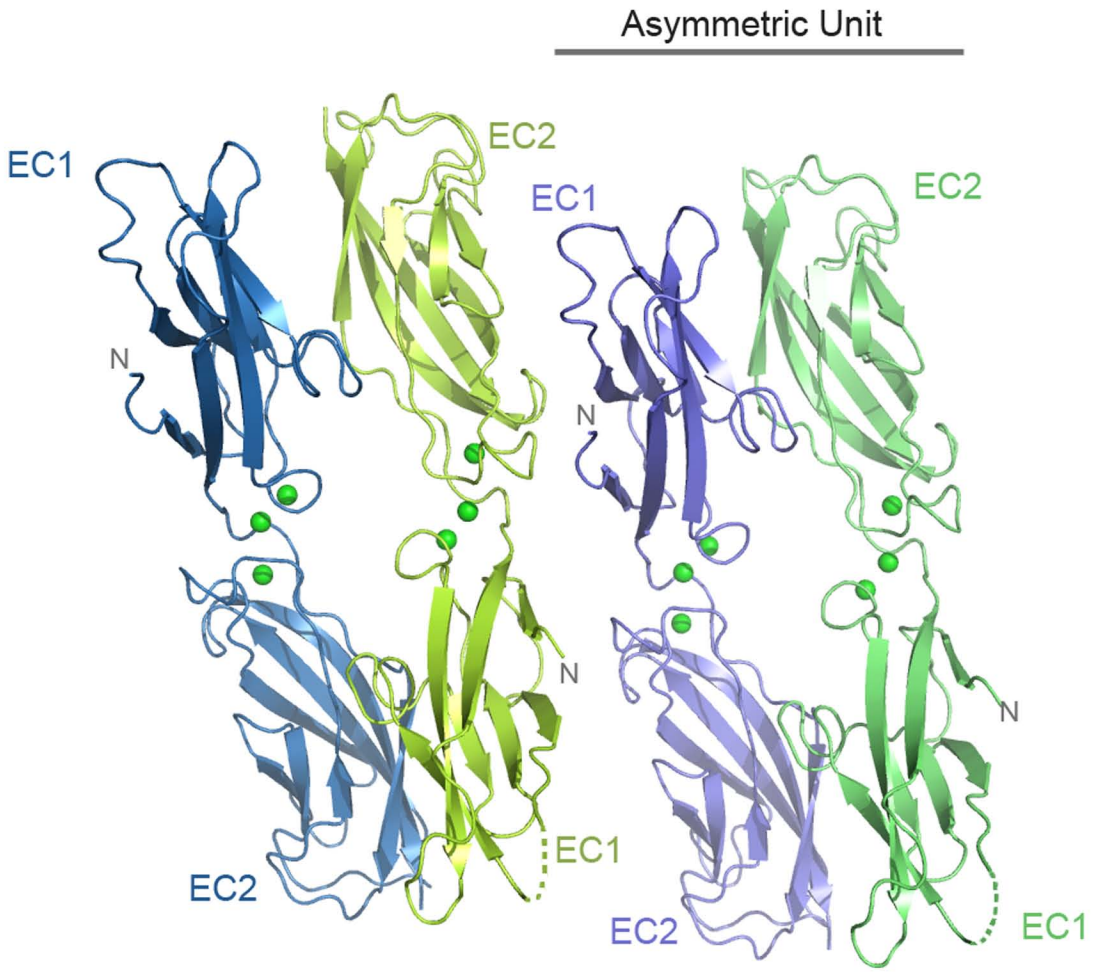
Supplementary Figure 3: Comparison of residues forming the X-interface in T-cadherin, type I and type II classical cadherins. The alignment is derived from alignment of the X-dimer crystal structures chicken T-cadherin, mouse E-cadherin and mouse cadherin-6 (colored blue) and from the sequence alignment of additional sequences within each cadherin subgroup (*i.e.* T-cadherins, type I and type II cadherins respectively). Within each subgroup, the X-interface residues strictly conserved are highlighted in dark blue and those conserved in character are highlighted in dark green. Two pairs of residues forming salt bridges in T-cadherin and E-cadherin crystal structures are indicated by, respectively, 1, 1' and 2, 2'. It is remarkable that the residues that form the second salt bridge in the mouse E-cadherin X-dimer (Arg105 and Asp199) are correlatively mutated (to Asp105 and Lys-200) in *Xenopus* E-cadherin. Secondary structure elements are indicated by yellow arrows above the alignments.



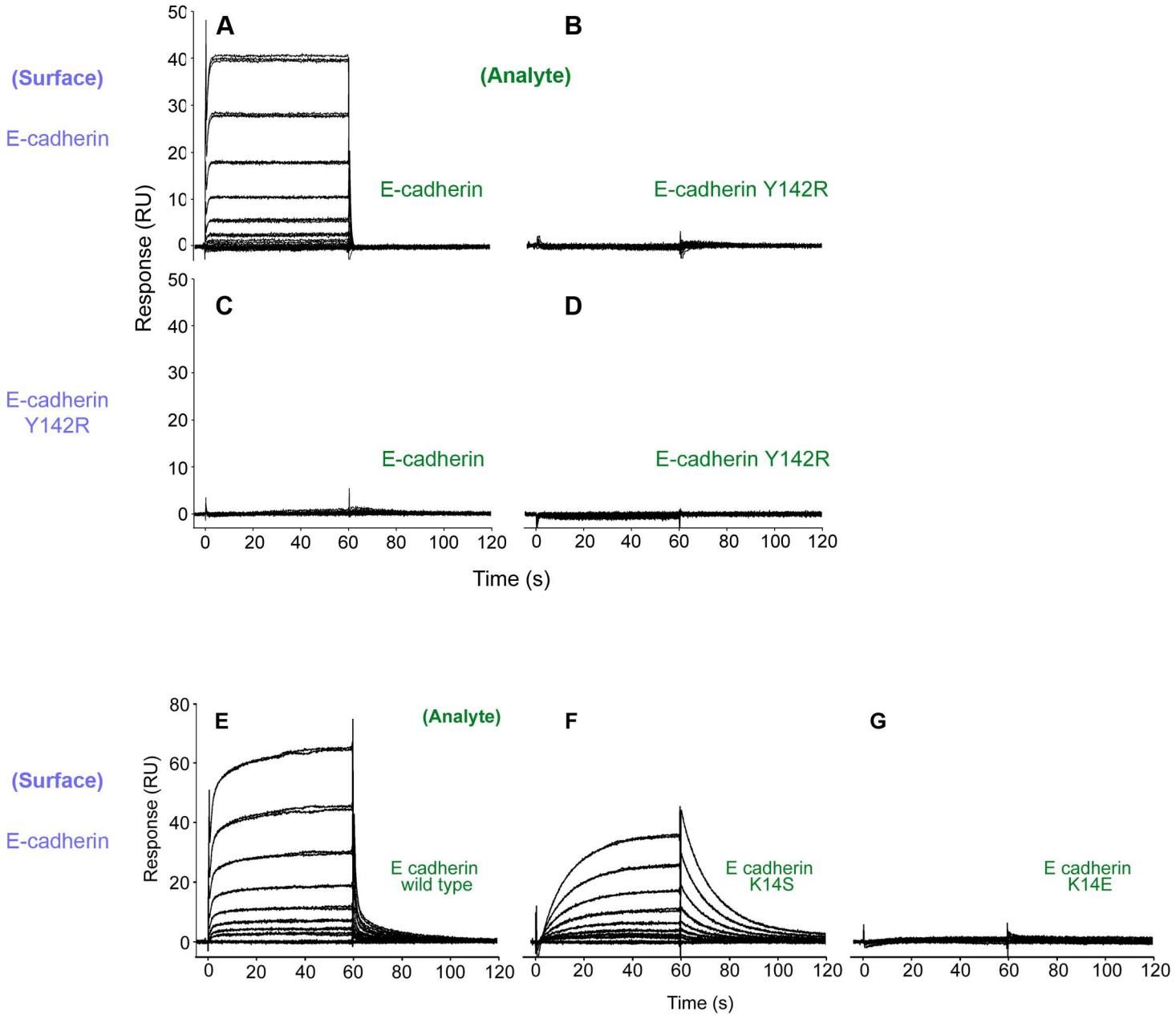
Supplementary Figure 4: Buried surface area by residue for X-dimer interfaces in T-cadherin, mutant E-cadherin and mutant cadherin-6. Buried surface area (BSA, in Å^2) is plotted for residues involved in the X-dimer interfaces of T-cadherin, E-cadherin W2A[2] and cadherin-6 W4A. Residues that participate in salt bridges in the dimer are boxed. Each protomer of the cadherin-6 X-dimer is plotted separately, given the slight asymmetry of this dimer (see main text). Asterisks indicate positions with significant differences in hydrophobic contacts.



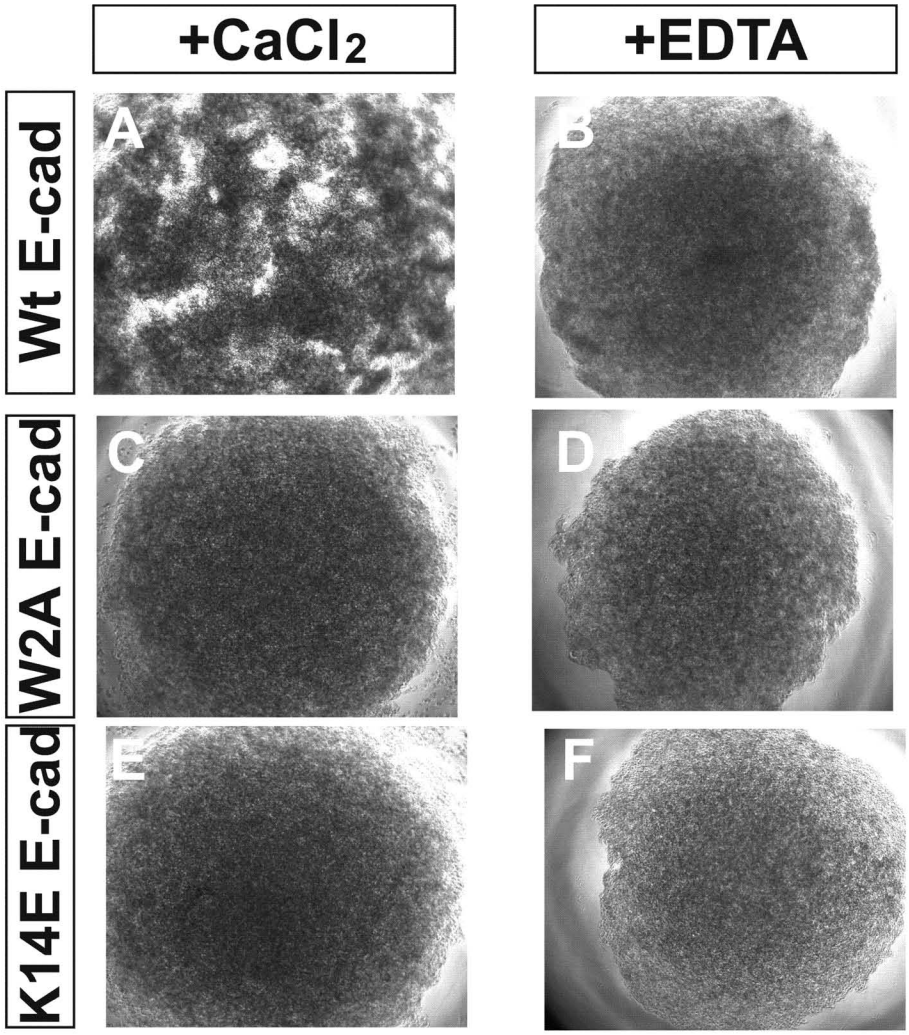
Supplementary Figure 5: Sedimentation velocity AUC profiles of wild-type and K14E E-cadherin EC1-2 fragments at 43µM protein concentration. (A) Wild-type E-cadherin; (B) K14E mutant. These experiments are identical to those shown Fig. 5, but are run using a lower protein concentration (see main text).



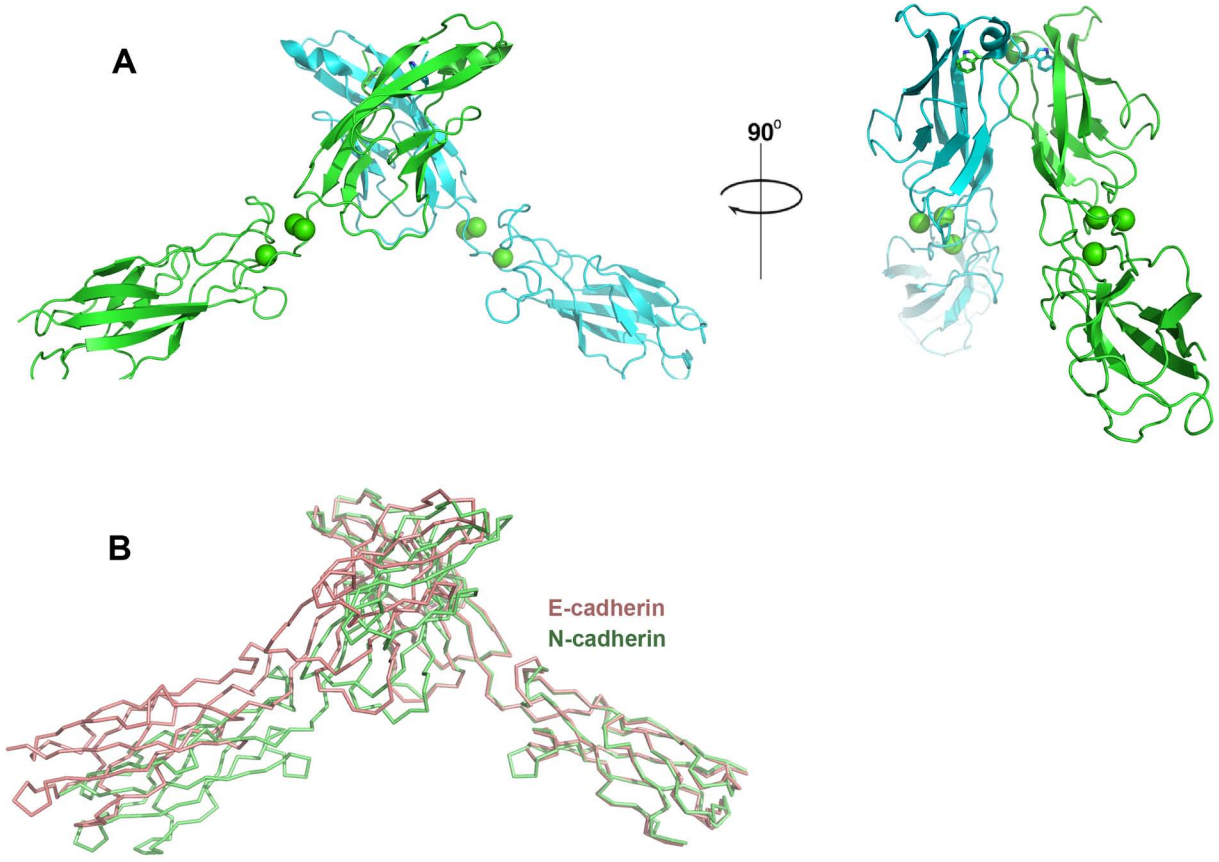
Supplementary Figure 6: Structure of E-cadherin EC1-2 W2A K14E double mutant
Ribbon view of a region of the crystal lattice comprising four cadherin monomers in crystals of E-cadherin EC1-2 W2A K14E mutant in which both strand swapping and X-dimerization were compromised by point mutations. No strand dimerization or X-dimer formation was observed. Residues 1 through 4 of the A-strand were disordered (marked ´N´). Calcium ions are shown as green spheres; disordered regions (the BC loop of EC1) shown as dashed lines.



Supplementary Figure 7: SPR binding analysis of E-cadherin Y142R and K14S X-dimer mutants. Immobilized wild-type EC1-2 protein supported substantial homophilic binding of wild-type EC1-2 E-cadherin (A), however, no binding of the EC1-2 Y142R mutant protein was detected (B). Similarly, immobilized Y142R supported no detectable binding of wild-type (C) or Y142R mutant protein (D). In a separate experiment, immobilized wild-type EC1-2 E-cadherin supported binding of wild-type protein (E) and a diminished level of binding of K14S mutant EC1-2 E-cadherin (F). As observed previously, K14E showed no detectable binding to the wild-type E-cadherin surface (G). All analytes were tested at a concentration range of 100-0.39μM, in duplicate. KD values from AUC analysis are shown in Table 2.



Supplementary Figure 8: Hanging drop aggregation assays for E-cadherin, strand dimer and X-dimer mutant CHO cells. (A) Wild-type E-cadherin expressing CHO cells produce cell aggregates that spread throughout the drop. However, both E-cadherin W2A (C) and E-cadherin K14E (E) expressing CHO cells form a thin sheet of cells at the meniscus of the drop with no aggregation observed. This configuration is similar to that found in the presence of 4mM EDTA for E-cadherin (B), E-cadherin W2A (D) and E-cadherin K14E (F).



Supplementary Figure 9: Crystal structure of N-cadherin EC1-2. (A) Strand swapped dimer formed between symmetry related protomers in crystals of mouse N-cadherin EC1-2. Two ribbon views, rotated 90°, are shown. The side chains of Trp2 residues in the swapped strands are displayed; calcium ions are shown as green spheres. (B) Superposition of the alpha carbon trace of the strand dimer formed by N-cadherin EC1-2 (red) with that formed by E-cadherin EC1-2 (green). Superposition based on one protomer (shown in background).

	Type of Interface	Total BSA (Å^2)	Buried Surface Area by domain (Å^2)		
			EC1	Linker region	EC2
T-cadherin WT	X-dimer	2173.0	926.0	447.6	799.4
E-cadherin WT (2O72.pdb)	Swapped	1800.4	1800.4	0	0
E-cadherin W2A	X-dimer	2209.1	1013.7	386.8	808.6
E-cadherin E89A	X-dimer	2176.9	1010.1	390.2	776.6
E-cadherin AA N-term	X-dimer	2258.2	1046.1	375.8	836.3
E-cadherin MR N-term. (1EDH)	X-dimer	2170.6	1001.2	383.5	785.9
E-cadherin M N-term. (1FF5)	X-dimer	2271.2	1132.5	373.0	765.7
Cadherin-11 WT (2A4E)	Swapped	2696.0	2696.0	0	0
Cadherin-6 W4A	X-dimer [1]	2692.3	1277.8	309.8	1104.7
		2496.0	1080.1	302.6	1113.3

1. BSA values for two dimers in crystallographic asymmetric unit are tabulated separately

Supplementary Table 1: Buried surface areas (BSA) of strand-swapped and X-dimer interfaces of type I and type II classical cadherin.

Supplementary Text

Comparison of X-dimer interfaces of E-cadherin, T-cadherin and cadherin-6

X-dimers of T-cadherin [1], mutant cadherin-6 (this work) and mutant E-cadherin (this work, [2,3]) can be superposed with rms deviations in Cα atoms of less than 3Å over at least 355 residues (Supplementary Fig. 1), and 2.6 Å or less when the comparison is restricted to the 44-46 residues comprising the X-dimer interfaces. Given the significant divergence in primary sequence between these cadherins (cadherin-6 and T-cadherin share 35% and 40% identity with E-cadherin over EC1-2 in mouse), this is a remarkable finding. Although the dimer interactions involve identical regions of the proteins, specific interactions at the atomic level are distinct. We compared three selected regions which together encompass the X-interfaces of T-cadherin, mutant E-cadherin and mutant cadherin-6 (Supplementary Fig. 2).

In the first region (Supplementary Fig. 2A, Region 1) the AB loop containing residue 14 in EC1 contacts the BC loop comprising residues 136-140 in EC2 of the partner protomer. T- and E-cadherin (Supplementary Fig. 2B, and C) have a basic residue at position 14 (Arg14 or Lys14 respectively) which forms a salt bridge with an acidic residue in the EC2 BC loop (Asp140 or Asp138). In contrast, cadherin-6 (Supplementary Fig. 2D) has an uncharged threonine at this position so no equivalent salt bridge is formed. Instead, Thr14 forms hydrogen bonds with backbone atoms of Pro136 and Thr137.

Similarly, cadherin-6 differs from T- and E-cadherin in the EC1-2 linker (Region 2 in Supplementary Fig. 2A). T- and E-cadherin (Supplementary Fig. 2E, F) have equivalent polar residues (Gln100 and Gln101 respectively) that form direct intermolecular hydrogen bonds with a conserved asparagine residue in EC2 (Asn141 or Asn143) and backbone atoms of Asp99 or Asp100. Cadherin-6 has a non-polar residue in the equivalent position (Ile99), so no analogous hydrogen bonds are formed (Supplementary Fig. 2G). Packing of the non-polar isoleucine side chains may instead contribute to dimer stability. The E-cadherin (W2A) X-dimer also has a network of ordered water molecules that mediate indirect hydrogen bonding between opposing protomers in the linker region,

as was observed in published E-cadherin X-dimer structures [2,3] (Supplementary Fig. 2F). A similar network may also exist in the T-cadherin X-dimer given the conservation of several relevant residues (Gln100, Asn141, Asn101). However, in the T-cadherin structure determined at 2.8Å resolution, water molecules were not included in the structural model due to a lack of clear water electron density [1].

In the region of the dimer where the FG loops at the apex of EC2 make intermolecular contact (Supplementary Fig. 2, Region 3), T- and E-cadherin but not cadherin-6 form equivalent intermolecular salt bridges between Arg104 and Asp202 or Arg105 and Glu199 respectively (Supplementary Fig. 2H, I, and J). However, all three structures have non-polar residues in the FG loop of EC2 whose side chains project into the interior of the dimer and pack together. In T-cadherin these are Met198 and Leu205; in E-cadherin Leu196 and Leu201 and in cadherin-6 Met188, Met192 and Leu195 (Supplementary Fig. 2H, I, and J; Supplementary Fig. 4).

In addition to the contacts highlighted in Figure S2, the A strands of EC1 are apposed in the X-dimer. Pro5, Pro6 and Pro10 residues in E-cadherin; Met3, Phe8 and Leu10 residues in cadherin-6 and Leu8 and Pro10 residues in T-cadherin contribute hydrophobic buried surface area to the interface and may be important for its stability.

Sequence alignment of type I, type II and T-cadherin EC1-2 domains confirms that the positions of residues involved in the three types of X-dimer interface are topologically equivalent and shows conservation of some components of the interface in related cadherins (Supplementary Figs. 3 and 4). Hydrophobic residues at the cadherin-6 X-interface are highly conserved among type II cadherins, and a majority of the T-cadherin X-interface residues, including those involved in the two salt-bridges, are conserved between T-cadherin species. Among type I cadherins, some variations exist in the composition and the size of the contacting FG loop region, and the residues forming the two salt-bridges in the E-cadherin X-dimer are not strictly conserved (Supplementary Fig. 3, compare E- and N-cadherin). It is remarkable that the residues which form the second salt-bridge in mouse E-cadherin (Arg105 and Glu-199) are correlatively mutated to

Asp105 and Lys200 in *Xenopus* E-cadherin. This indicates that selective pressure may have maintained this salt-bridge and supports the possibility of a functional role for it (Supplementary Fig. 3).

Together, structural and sequence comparisons suggest that although the X-dimer interface has been preserved in type I, type II, and T-cadherins, the atomic-level interactions have varied. In particular, the interface in the type II cadherin appears divergent compared with that of T- and E-cadherins.

Supplementary methods

Protein production

We cloned the coding sequences for the first two EC domains of mature E-cadherin (Asp1 to Asp 213), N-cadherin (Asp1 to Val216) and cadherin-6 (Ser1 to Asp207) from mouse into the BamH1 and Not1 sites of the prokaryotic expression vector pSMT3. This vector encodes an N-terminal 6xHis-SUMO protein tag followed by the inserted protein. Fusion protein cleavage with Ulp1 (Ubiquitin-like protease 1) after a di-glycine motif preceding the protein of interest allows isolation of untagged fragments with precise N-termini [4]. We removed bases encoding additional residues between the protease site and the cadherin N-terminus introduced by cloning by Quikchange site-directed mutagenesis (Stratagene). Thus, all cadherin EC1-2 proteins used in this study had native mature N-termini, except where specifically altered (E-cadherin AA-extension and DW-deletion mutants). We prepared constructs encoding point mutants of E-cadherin EC1-2 and cadherin-6 EC1-2 by Quikchange site-directed mutagenesis (Stratagene)

We expressed proteins in *E.coli* Rosetta 2 (DE3) pLysS cells (Novagen). We grew cultures at 37°C to an OD_{600} of 0.6 then induced with 0.1mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) and transferred to 18°C for 16 hours (E- and N-cadherin) or to 30°C for 5 hours (cadherin-6). We harvested cells by centrifugation (6000g, 15min) and resuspended in binding buffer (500mM NaCl, 20mM Tris-Cl pH 8.0, 3mM $CaCl_2$, 20mM imidazole) before sonication for 3 minutes at 4°C. Cell debris was pelleted by centrifugation (20000g, 1 hour). We purified 6xHis-SUMO-cadherin fusion proteins from the lysate by affinity chromatography using IMAC Sepharose 6 Fast Flow resin (GE Healthcare) charged with Ni^{2+}. After washing with binding buffer, we eluted proteins with 250mM imidazole. The 6xHis-SUMO tag was removed by 0.1mg of 6xHis tagged Ulp1 protease during overnight dialysis at 4°C into 50mM NaCl, 20mM Tris-Cl pH 8.0, 3mM $CaCl_2$. We removed cleaved fusion tag, uncleaved fusion protein and protease in a second nickel affinity chromatography step and purified cadherin EC1-2 proteins in the flow-through further by ion exchange and size-exclusion chromatography. We concentrated purified proteins to approximately 8mg mL^{-1} in 150mM NaCl, 10mM Tris-

Cl pH 8.0, 3mM Ca Cl_2, except for N-cadherin EC1-2 which was concentrated to 2mg mL^{-1} owing to poor solubility.

We prepared cadherin EC1-2 proteins with specifically biotinylated C-termini for SPR studies using the AviTag system (Avidity), by including a sequence encoding an 'AviTag' (GGGLNDIFEAQKIEWHE) before the stop codon. *E. coli* BL21 (DE3) cells (Novagen) were co transformed with the cadherin constructs and an IPTG-inducible expression plasmid encoding BirA biotin ligase and grown to an OD_{600} of 0.6 before 0.1mM IPTG and 50μM biotin were added to induce expression and allow specific biotinylation of the AviTag for 4 hours at 37°C. We carried the purification of tagged proteins out exactly as above. We confirmed biotinylation by Western blot with avidin-horseradish peroxidase (Sigma).

Analytical size-exclusion chromatography

We ran purified cadherin EC1-2 fragments over a Superose 12 10/30 column (GE Healthcare) equilibrated with running buffer (150mM NaCl, 10mM Tris-Cl pH 8.0, 3mM $CaCl_2$). For E-cadherin, we included 0.3mM TCEP in the buffer (identical results were obtained in the absence of reductant). We diluted proteins to the desired concentration in running buffer 4 hours before the experiment and stored at 4°C. 150μL of protein sample was injected onto the column and eluted with binding buffer at a flow rate of 0.3mL min^{-1} at 4°C on an Akta FPLC system (GE Healthcare). Total time for the separation was 30 to 40 minutes. Elution of protein was monitored by absorbance at 280nm and fractions were collected for SDS PAGE analysis. We determined the void volume of the column to be approximately 7.7mL using Blue Dextran (GE healthcare).

CHO cell line production

A DNA construct encoding full-length mouse E-cadherin in pcDNA 3.1 (Invitrogen) was kindly provided by Dr. N. Zampieri (Dept. of Neurobiology, Columbia University). We introduced K14E and W2A point mutations into this construct using the Quikchange method (Stratagene). We transfected constructs into CHO cells using Lipofectamine 2000 (Invitrogen) following the manufacturers protocol. Following four weeks selection

using G418 sulfate (600μg mL^{-1}) individual clones of CHO cells were isolated and cultured on glass coverslips to assess E-cadherin expression. We selected clones expressing similar levels of cadherin and expanded these for aggregation experiments.

Aggregation assays

A complete dispersed cell suspension was obtained by treating confluent CHO cells with enzyme free cell dissociation solution (Sigma) at 37°C for 20 minutes. We performed long-term aggregation assays essentially as Katsamba et al., 2009. We resuspended the Cells in Dulbecco's Modified Eagle Media (DMEM) containing 10% (v/v) Fetal Calf serum (Gibco) and 70 units of DNAse 1 (Sigma). $5x10^4$ cells per 0.5mL were added to 24 well Ultra low cluster plates (Corning Costar, Cambridge, MA U.S.A) and allowed to aggregate for 24hrs at 37°C on a rotary shaker with 70-80 revolutions per min in a humidified atmosphere comprising 5% $C0_2$/95%air. To remove Ca^{2+} for control experiments, we added ethylenediaminetetraacetic acid (EDTA, pH 8.0) to a final concentration of 4mM. The parental CHO cell line showed no aggregation in these conditions. We performed hanging drop assays as in Katsamba et al. (2009). Following labelling and dissociation, $5x10^3$ cells of each type were placed in a 25μL drop on the lid of a 10cm petri dish above a reservoir of 10mL of culture medium. After 16 hours at 37°C in a humidified atmosphere comprising 5% $C0_2$/ 95%air the wells were imaged using a Nikon DS5M camera attached to a Nikon TS100 inverted microscope using a 4x objective lens.

Antibody staining

We grew CHO cells expressing E-cadherin on acid washed coverslips until defined clusters of cells had formed. We then fixed cells in 2% (w/v) Paraformaldehyde for 20 minutes at 4°C, washed these three times in phosphate buffered saline (PBS), and incubated these in blocking solution (PBS/10% (v/v) FCS/ 0.1% (v/v) Triton X-100) for 30 minutes at 20°C. We added rat anti-E-cadherin (clone ECC, Invitrogen) diluted 1:1000 in blocking solution for 16 hours at 4°C. Cadherin localization in non-permeabilized conditions omitted Triton X-100 from the protocol. A CY3- conjugated Goat anti Rat secondary antibody was used (Jackson immunoresearch). We mounted

coverslips in Vectashield (Vector Laboratories) and imaged these under 20x objective magnification on a Nikon E80i microscope fitted with Hammamatsu ORCA-ER digital camera. Controls of native CHO cells or E-cadherin cells incubated without primary antibody showed no significant staining using this protocol.

Supplementary references

1. Ciatto, C. et al. T-cadherin structures reveal a novel adhesive binding mechanism *Nature Structural and Molecular Biology*, In Press (2009).
2. Nagar, B., Overduin, M., Ikura, M. & Rini, J.M. Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature* **380**, 360-4 (1996).
3. Pertz, O. et al. A new crystal structure, Ca2+ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. *EMBO J* **18**, 1738-47 (1999).
4. Mossessova, E. & Lima, C.D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* **5**, 865-76 (2000).